

zation 3:79-94.

Phillips, A.T., J.B. Rosen, and M. van Vliet. 1992. A parallel stochastic method for solving linearly constrained concave global minimization problems. *Journal of Global Optimization* 2:243-258.

Rinnooy Kan, A.H.G. and G.T. Timmer. 1987. Stochastic global optimization methods. Part I: Clustering methods. *Mathematical Programming* 39(1):27-56.

Tuy, H. 1964. Concave programming under linear constraints. *Doklady Akademii Nauk SSSR* 159:32-35. English translation in *Soviet Mathematics Doklady* 5:1437-1440.

- Hansen, P., B. Jaumard, and S.H. Lu. 1991. An analytical approach to global optimization. *Mathematical Programming* 52(2):227-254.
- Horst, R. 1984. On the global minimization of concave functions: Introduction and survey. *Operations Research Spektrum* 6:195-205.
- McCormick, G.P. 1976. Computability of global solutions to factorable nonconvex programs: Part I - convex underestimating problems. *Mathematical Programming* 10(2):147-175.
- Mockus, J. 1989. Bayesian approach to global optimization. Dordrecht: Kluwer Academic Publishers.
- Pardalos, P.M., and C.A. Floudas. 1990. A collection of test problems for constrained global optimization algorithms. In *Lecture Notes in Computer Science* 455, ed. G. Goos and J. Hartmanis. Berlin: Springer-Verlag.
- Pardalos, P.M., and J.B. Rosen. 1986. Methods for global concave minimization: A bibliographic survey. *SIAM Review* 28(3):367-379.
- Pardalos, P.M., and J.B. Rosen. 1987. Constrained global optimization: Algorithms and applications. In *Lecture Notes in Computer Science* 268, ed. G. Goos and J. Hartmanis. Berlin: Springer-Verlag.
- Phillips, A.T. 1988. Parallel Algorithms for Constrained Optimization. Ph.D. dissertation, University of Minnesota, Minneapolis, MN.
- Phillips, A.T., and J.B. Rosen. 1988. A parallel algorithm for constrained concave quadratic global minimization. *Mathematical Programming* 42(2):421-448.
- Phillips, A.T., and J.B. Rosen. 1990. A parallel algorithm for partially separable non-convex global minimization: Linear Constraints. *Annals of Operations Research* 25:101-118.
- Phillips, A.T., and J.B. Rosen. 1992. Sufficient conditions for fast solution of linearly constrained global minimization problems. *Journal of Global Optimi-*

(Averages for 10 problems)

Domain		Sufficient Conditions		Subregion	Incumbent	
<u>mn</u>	<u>p</u>	<u>Splits</u>	<u>Satisfied</u>	<u>Eliminations</u>	<u>Improvements</u>	
10	10	0	5.3	1.0	0.9	4.8
10	10	10	2.7	0.9	0.3	0.2
10	10	20	2.5	0.5	0.4	20.2
10	10	30	4.4	0.5	0.6	1.8
10	10	40	3.2	0.3	0.6	1.1
10	20	0	14.2	1.1	1.7	1.5
10	20	10	7.3	0.9	0.9	4.1
10	20	20	2.7	0.7	1.0	3.4
10	20	40	6.5	0.7	0.9	2.1
10	30	0	6.4	1.0	0.8	5.6
10	30	10	17.5	1.0	1.5	7.6
10	30	20	2.7	0.4	1.1	1.5
10	40	0	9.3	0.9	0.8	4.1
10	40	10	7.8	0.7	1.6	12.1

5 Acknowledgments

This research was supported in part by the grant AFOSR-91-0147 and by the Army High Performance Computing Research Center and the Minnesota Supercomputer Institute.

6 References

- Boender, C.G.E., and A.H.G. Rinnooy Kan. 1987. Bayesian stopping rules for global optimization methods. *Mathematical Programming* 37(1):59-80.
- Falk, J.E., and K.R. Hoffman. 1976. A successive underestimation method for concave minimization problems. *Mathematics of Operations Research* 1(3):251-259.
- Falk, J.E., and R.M. Soland. 1969. An algorithm for separable nonconvex programming problems. *Management Science* 15(9):550-569.

branch and bound aspect of the method. In addition, Table 5 lists the average number of times the sufficient conditions were satisfied, the average number of times that a subregion was eliminated by the heuristic linear underestimation step, and the average number of times that the incumbent solution was improved.

Finally, it should be noted that both algorithms have been useful in obtaining solutions to problems which are better than the previously reported “global” solutions. In particular, the collection of global optimization of very difficult test problems compiled by Pardalos and Floudas (1990) contains a number of problems of the form (GP). The solutions reported in problems 2.1 through 2.7.4 have all been confirmed by both the stochastic and deterministic approaches. However, Pardalos and Floudas (1990) report a “best known solution” of $\phi = -4105.28$ with corresponding vertex x (non-zero components only) of $x_4 = 0.995$, $x_{12} = 0.930$, $x_{16} = 7.412$, $x_{18} = 12.674$, and $x_{20} = 17.990$ for problem 2.7.5, whereas both the stochastic and deterministic methods discovered the global minimum function value $\phi^* = -4150.41$ with corresponding vertex x^* (non-zero components only) of $x_3 = 1.043$, $x_{11} = 1.747$, $x_{13} = 0.431$, $x_{16} = 4.433$, $x_{18} = 15.859$, and $x_{20} = 16.487$.

Table 5

Deterministic Method Statistics

(Averages for 10 problems)

<u>mn</u>	<u>p</u>	Global MinTrials Since Last				
		<u>Trials</u>	<u>Local Minima</u>	<u>was Local Min #</u>	<u>Local Min Found</u>	
10	10	0	261.0	25.0	4.9	41.8
10	10	10	247.0	23.6	4.9	35.3
10	10	20	219.0	20.8	6.4	31.0
10	10	30	190.0	17.9	6.5	40.9
10	10	40	159.0	14.8	3.7	40.8
10	20	0	992.0	98.1	5.2	38.6
10	20	10	874.0	86.3	8.9	44.5
10	20	20	671.0	66.0	7.4	41.9
10	20	40	505.0	49.4	8.7	41.3
10	30	0	2245.0	223.4	13.4	49.7
10	30	10	1808.0	179.7	22.8	38.1
10	30	20	1119.0	110.8	6.9	26.4
10	40	0	3937.0	392.6	8.1	48.2
10	40	10	2596.6	271.4	11.9	23.5

As described in section 2 above, the use of the bayesian stopping rule involving only the number of local minima and the number of trials is very often too conservative, and an additional test that allows termination of the method only after 99% of the feasible region has been explored would be more practical. In fact, these computational results fully support this hypothesis since in *all* 140 problems tested, the stochastic method was stopped by this additional check on the fraction of the domain explored. This greatly decreases the overall solution time since for a problem of size $m = 10$, $n = 20$, and $p = 0$, the total number of local minima discovered is approximately 98 and the original bayesian stopping rule alone would have required 19504 trials for termination. The results from Table 4 indicate that only 992 trials were needed using the combination of the two stopping rules.

Table 5 presents a more detailed set of statistics for the deterministic method. For each set of problems of the same size, this table lists the average number of domain splits required (in order to guarantee a global solution) by the

Table 3

Stochastic Method				Deterministic Method			
<u>mn</u>	<u>p</u>	<u>Time (secs)LPs</u>		<u>Pivots/LPTime (secs)LPs</u>		<u>Pivots/LP</u>	
10	10	0	0.73	841.2	3.0	0.43	470.82.8
10	10	10	1.80	904.7	5.3	0.34	177.33.9
10	10	20	2.74	768.3	7.4	0.98	585.63.8
10	10	30	3.49	653.5	8.7	0.97	306.23.8
10	10	40	3.99	579.2	9.8	1.12	260.74.1
10	20	0	6.16	3719.5	4.3	6.59	5885.42.0
10	20	10	10.16	3370.7	6.0	68.83	79671.11.9
10	20	20	12.14	2739.3	7.1	1.13	524.72.9
10	20	40	15.19	1985.3	9.2	3.02	984.42.3
10	30	0	23.58	8640.1	5.3	2.16	1622.72.3
10	30	10	29.53	7619.0	6.3	5.80	3428.12.7
10	30	20	24.90	4599.1	7.3	2.12	937.22.8
10	40	0	62.76	15997.5	6.0	4.01	2418.71.4
10	40	10	54.50	10532.8	6.9	6.66	3960.51.4

Table 4 presents a more detailed set of statistics for the stochastic method.

For each set of problems of the same size, this table lists the average number of random search directions required (i.e. trials) and the average number of local minima found. In addition, since the global minimum vertex is one of the local minima, Table 4 also shows, on average, which local minimum turned out to be the global one. From this table it is apparent that the global minimum is detected relatively early in the process, but the remaining trials are still required in order to satisfy the bayesian stopping rules. Finally, once the final local minimum vertex is found, the stochastic method must still perform a number of local searches in order to satisfy the stopping criteria. The number of such trials is listed in the last column of Table 4.

Table 4

Stochastic Method Statistics

Time -- Stochastic					Time -- Deterministic		
(seconds)					(seconds)		
<u>mn</u>	<u>p</u>	<u>min</u>	<u>max</u>	<u>avg</u>	<u>min</u>	<u>max</u>	<u>avg</u>
10	10	0	0.42	1.60	0.73	0.05	2.520.43
10	10	10	1.19	2.65	1.80	0.08	0.630.34
10	10	20	1.85	3.78	2.74	0.20	5.380.98
10	10	30	2.74	4.64	3.49	0.42	2.750.97
10	10	40	2.44	5.79	3.99	0.42	3.341.12
10	20	0	2.90	13.63	6.16	0.28	33.916.59
10	20	10	4.51	16.17	10.16	0.33	675.0068.83
10	20	20	7.16	19.05	12.14	0.39	2.251.13
10	20	40	7.51	25.13	15.19	1.21	4.973.02
10	30	0	10.39	40.29	23.58	0.32	5.952.16
10	30	10	17.81	59.17	29.53	0.42	23.525.80
10	30	20	9.48	46.06	24.90	0.74	7.792.12
10	40	0	16.67	99.14	62.76	1.15	9.934.01
10	40	10	18.69	99.86	54.50	1.14	21.636.66

It is clear from this table that in almost every case the deterministic method solves fewer linear programs, requires fewer pivots per linear program, and takes substantially less overall time to obtain a global solution than does the stochastic method. In two cases ($n = 20$ and $p = 0$; $n = 30$ and $p = 10$), the branch and bound portion of the deterministic method had to perform an unusually large number of domain splits in order to guarantee a global solution; hence, for these two cases the number of linear programs solved and the overall solution time are much larger than would otherwise be expected. In fact, for $n = 20$ and $p = 0$, the stochastic method turned out to be faster, on average, than the deterministic method. Furthermore, the results cited for the deterministic method in the case $n = 20$ and $p = 10$ were almost entirely dominated by a single very difficult problem which required 675 seconds and 788467 linear programs to solve. Excluding this particular problem, the average overall solution time for a problem of this size would have been 1.47 seconds for the deterministic method.

set of 10 problems. These results indicate, as stated above, that the addition of purely linear variables tends to make the problems somewhat easier as p begins to dominate n .

Table 1

Number of Local Minima

<u>m</u>	<u>n</u>	<u>p</u>	<u>min</u>	<u>max</u>	<u>avg</u>
10	10	0	13	57	25
10	10	10	18	33	24
10	10	20	16	28	21
10	10	30	13	22	18
10	10	40	9	23	15
10	20	0	48	217	98
10	20	10	44	141	86
10	20	20	43	103	66
10	20	40	24	83	49
10	30	0	106	381	223
10	30	10	108	356	180
10	30	20	44	200	111
10	40	0	116	587	393
10	40	10	87	600	271

Table 2 displays the minimum, maximum, and average CPU solution time required for each algorithm on the same set of problems. Table 3 presents averages of the total time required (repeated from table 2), the total number of linear programs solved, and the average number of pivots per linear program. In 138 out of the 140 problems tested, the two methods obtained exactly the same global minimum vertex. For the two cases in which the stochastic method failed to find the global minimum vertex, the vertex that it found differed in function value from the true global minimum function value (as given by the deterministic method) in each case by less than 2.5% (the exact relative errors were 2.5% and 0.06%).

Table 2

the global minimum to within any specified tolerance in a finite number of iterations.

4 Computational Results

Computational results reported in Phillips, Rosen, and van Vliet (1992) indicate that randomly generated concave quadratic problems for which the global *maximum* is known to be interior to the polytope are among the most difficult test problems available. Hence, the class of separable test problems of the form (GP) used for comparing the two methods were randomly generated concave quadratic functions of the following form:

where v is the unconstrained global maximum of $\phi(x)$, and $\lambda_i < 0$ for $i=1, \dots, n_1$. As stated above, it has been observed (Phillips, Rosen, and van Vliet 1992) that for problems with the global maximum $v \in _$, there exist significantly more local minima than for problems of the same dimension with no restriction on v , hence these problems should be more difficult to solve. For this reason, all problems were generated in such a way that $v \in _$. In addition, a purely linear term of the form $d^t y$, where $d, y \in \mathbf{R}^p$, was also added to the function $\phi(x)$. Both algorithms can be easily extended to handle these purely linear variables, and their presence is expected to make the problems somewhat easier when p dominates n (see table 1 below).

Over 140 test problems were used in the computational comparisons, and all results were obtained on the Cray x-mp ea/464 supercomputer located at the Minnesota Supercomputer Center in Minneapolis, MN. For each of the methods, the same set of 10 problems with dimensions of $m = 10$, $n \in \{10, 20, 30, 40\}$, and $p \in \{0, 10, 20, 30, 40\}$ were tested. Table 1 displays the minimum, maximum, and average number of local minima, as determined by the stochastic method, for each

If the difference $\varphi(x') - \Gamma(x')$ exceeds ε and the optimality check fails, then the following heuristic steps of the algorithm are performed. The feasible region is divided into $2n$ subregions R_{lj} (for $l=1, \dots, n$ and $j=1,2$) by bisecting R in certain specified directions. Linear underestimating functions $\Gamma_{lj}(x)$ to $\varphi(x)$ over the *subregions* are then constructed and minimized over the original region $_$. If the optimal function values for the linear problems are denoted by Γ_{lj} and the corresponding optimal solution points by x_{lj} , and if $\Gamma_{lj} > \varphi(x_{lj})$ for either $j = 1$ or 2 , then R_{lj} can be *eliminated* from further consideration. In addition to this domain reduction process, these linear underestimating problems also provide improved upper and lower bounds on the global minimum function value. Furthermore, the sufficient conditions test can again be performed and the ε tolerance checked. If both termination tests fail, and as long as some subregion of the original hyperrectangle R is eliminated at each step, the linear underestimating process just described is repeated on the smaller subregion.

If at some step no subregion can be eliminated and the current incumbent solution vertex does not satisfy the sufficient conditions or the ε tolerance, then the remaining hyperrectangle R is bisected and the entire procedure is repeated on the smaller hyperrectangles. This part of the method is represents the branch and bound aspect of the algorithm and provides for guaranteed termination at the global minimum solution. As before, for a more detailed discussion and the theoretical justifications of each of the above steps, including the sufficient conditions theory, see Phillips and Rosen (1992).

The major drawback of this deterministic approach is that separability of the objective function is required in order to make proper use of the sufficient conditions test. In addition, this approach does not produce any *local* minima except for the final global minimum. The main advantages of this approach are that it provides bounds on the global minimum at every step, and it is guaranteed to find

disadvantage to this approach is that it does not provide bounds on the global minimum, nor is it guaranteed to find the global minimum.

3 The Deterministic Approach

The deterministic approach uses a combination of linear underestimating subproblems, branch and bound techniques, and sufficient condition tests in order to provide a guaranteed global solution to the problem (GP). Since the feasible domain $_$ is assumed to be non-empty and bounded, the first step of this method is to construct a rectangular domain R which contains $_$. And since the concave function $\phi(x)$ is assumed to be separable, a linear function $\Gamma(x)$ which underestimates $\phi(x)$ over R and which agrees with $\phi(x)$ at *every vertex* of R can also be easily constructed. Hence, the solution to the linear underestimating subproblem

provides a feasible point x' of $_$ such that $\Gamma(x') \leq \phi^* \leq \phi(x')$. Thus an initial pair of lower and upper bounds is readily available.

The vertex x' is also a candidate for the global minimum, and if it happens that $\phi(x') - \Gamma(x') \leq \epsilon$ for some small user specified $\epsilon \geq 0$, then x' is usually accepted as the global optimum solution. In many of the previous computational methods (Phillips and Rosen 1988; Phillips and Rosen 1990), the solution x' obtained from this linear underestimating problem has, in fact, turned out to be the global optimum solution to problem (GP). Unfortunately, the difference $\phi(x') - \Gamma(x')$ usually exceeds ϵ for many iterations. These iterations are required only to verify, by improving the lower bound, that the point x' is indeed the global optimum solution. Hence, the deterministic method applies a sufficient conditions test to determine if the candidate vertex x' is indeed the global minimum vertex. This optimality check is described in detail in Phillips and Rosen (1992).

solving the linear problem

until two successive solution vertices z_{j-1} and z_j do not differ. Such a vertex is guaranteed to be a Karush-Kuhn-Tucker point, and hence a candidate for the global minimum (Phillips, Rosen, and van Vliet, 1992). Since the global minimum is unknown, the objective of this multistart method is to find *all* of the existing local minima for the problem. Hence, this procedure is repeated once again with another random search direction.

Unfortunately the total number of local minima is also unknown. Thus, an optimal bayesian estimate of the number of local minima is used to terminate the method. This bayesian stopping rule (Boender and Rinnooy Kan 1987) indicates that, with very high probability, all of the local minima have been found; hence, the one with the lowest function value will be the global minimum. The use of this stopping rule alone, as suggested by Boender and Rinnooy Kan (1987) and verified by Phillips, Rosen, and van Vliet (1992), is very often too conservative and an additional test incorporating the fraction of the feasible region explored would be more practical. Thus, the final step allows termination of the method only when 99% of the feasible region has been explored. For a more detailed discussion and the theoretical justifications of each of the above steps, including the stopping criteria, see Phillips, Rosen, and van Vliet (1992).

It should be noted that this stochastic algorithm can be applied to a much broader class of linearly constrained concave global minimization problems than are described in this paper. In fact, the concave function $\phi(x)$ need only be differentiable over \mathbb{R}^n for this approach to be applicable (separability is only required for the deterministic approach described in the next section). In addition, this approach produces all of the *local* minima in addition to the global minimum, which in many circumstances are as useful as the global solution itself. The main

The second method is a deterministic approach, described in detail in Phillips and Rosen (1992), which attempts to use a combination of linear underestimating subproblems, branch and bound techniques, and sufficient condition tests in order to recognize a solution to the global minimization problem (GP). In this method, a linear function which underestimates the original concave function is constructed, and the solution of the corresponding linear underestimating problem provides both upper and lower bounds on the global minimum function value. A heuristic step is applied in an attempt to eliminate parts of the feasible region which cannot contain the global minimum point. In the worst case, this step will fail to eliminate any regions, but typically the heuristic will allow the original feasible region to be rapidly reduced to a much smaller polytope in which the global minimum point must occur. Branch and bound techniques are then used to reduce the feasible region under consideration and decrease the difference between the upper and lower bounds. This procedure guarantees that an ϵ -approximate solution (the relative error in the objective function is bounded by a user specified tolerance ϵ) will be obtained. The use of sufficient conditions to recognize a global minimum, applied whenever a new candidate for the global minimum vertex is found, may significantly accelerate the solution for certain types of problems of the form (GP). If these sufficient conditions are not satisfied, the information can frequently be used to obtain improved bounds and possibly eliminate part of the feasible set from further consideration.

2 The Stochastic Approach

The stochastic approach consists of two main parts: a global and a local phase. In the global phase, a random search direction $u \in \mathbf{R}^n$ is selected and used to obtain a starting point from which the local phase may begin. The local phase then attempts to find a local minimum by starting from this point and repeatedly

ences to many of these applications. There has also been an active research effort on computational methods for solving problem (GP), and many of these methods are summarized in Pardalos and Rosen (1986) and Mockus (1989). Most methods are restricted to a certain class of concave functions, including quadratic functions (Phillips and Rosen 1988), non-convex separable functions (Falk and Soland 1969, Phillips and Rosen 1990), and factorable functions (McCormick 1976). Only a few methods have been developed for solving the general constrained concave minimization problem (Tuy 1964, Falk and Hoffman 1976, Horst 1984). Few of these methods, however, seem to be suitable for the efficient computational solution of problems with more than 25 variables. More recently, Hansen, Jaumard, and Lu (1991) have proposed a method for more general classes of nonconvex minimization problems which include concave problems as special cases.

This paper presents a computational comparison of two vastly different methods for solving problem (GP). The first of these methods is a stochastic approach, described in detail in Phillips, Rosen, and van Vliet (1992), and is based on a multistart technique first proposed by Rinnooy Kan and Timmer (1987). This technique repeatedly employs two phases during the solution process: a global and a local phase. In the global phase, a random search direction is selected and used to obtain a starting point from which the local phase may begin. The local phase then attempts to find a local minimum by starting from this point. Since the global minimum is generally unknown, the objective of a multistart method is to find all of the existing local minima for the problem. But, since the total number of local minima is also unknown, an optimal bayesian estimate of the number of local minima must be used to terminate the method. This bayesian stopping rule (Boender and Rinnooy Kan 1987) indicates that, with very high probability, all of the local minima have been found; hence, the one with the lowest function value will be the global minimum.

1 Introduction

This paper presents an in depth computational comparison of two methods for solving linearly constrained concave global minimization problems. In particular, the two methods presented are used to obtain a solutions to linearly constrained concave global minimization problems of the form

where $\varphi(x)$ can be expressed in separable form. That is,

and each $\varphi_i(x_i)$ is concave. Additionally, $_ = \{ x : Ax _ b, x _ 0 \}$ is assumed to be nonempty and bounded, and $x \in \mathbf{R}^n$, $A \in \mathbf{R}^{m \times n}$, and $b \in \mathbf{R}^m$.

Problem (GP) is a constrained combinatorial optimization problem for which many well known problems are special cases. For example, the concave quadratic global minimization problem is a special case of problem (GP) for which $\varphi(x) = (1/2)x^tQx + c^tx$ where $Q \in \mathbf{R}^{n \times n}$ is symmetric and negative definite ($\varphi(x)$ can be transformed into separable form using the eigenstructure of Q). This concave quadratic global minimization problem is known to be NP-hard (Phillips 1988), and hence it follows that problem (GP) is NP-hard. From a computational viewpoint, this means that, in the worst case, the computing time required to obtain a solution will grow exponentially with the number of nonlinear variables. An important property of problem (GP), which is basic to many solution methods (Phillips 1988), is that the global minimum point is always found at a vertex of the convex polytope $_$. For this reason, linear programming is an essential part of any computational algorithm to solve problem (GP).

A substantial literature exists describing applications of this global minimization problem. The monograph by Pardalos and Rosen (1987) contains refer-

A Computational Comparison of Two Methods for Constrained Global Optimization

A.T. Phillips^{1*}

J.B. Rosen^{2**}

abstract

For constrained concave global minimization problems, two very different solution techniques have been investigated. The first such method is a stochastic multistart approach which typically finds, with high probability, all local minima for the problem. The second method is deterministic and guarantees a global minimum solution to within any user specified tolerance. It is the purpose of this paper to make a careful comparison of these two methods on a range of test problems, and to investigate in this way the advantages and disadvantages of each method. A direct computational comparison, on the same set of over 140 problems, is presented.

Keywords: *global optimization, stochastic methods, deterministic methods.*

1. * Computer Science Department, United States Naval Academy, Annapolis, MD 21402.

2. ** Computer Science Department, University of Minnesota, Minneapolis, MN 55455.